

[illegible]

APPLICANT: ANTONIO CANTONI AND ROBERT M. NEWMAN

Express Mail Label No. EL224699579US

D.C. 20251.

July 31, 2001

Date of Deposit

Date of Deposit _____
Signature Samantha Bell

Signature Samantha Bell
Typed or Printed Name of Person Signing Certificate

This invention relates to a transfer of messages in a 5 multiplexed system.

More particularly, the invention relates to a method for the transfer of messages in a time multiplexed slotted environment such as a communications network. The network can be of the type disclosed in International Publication No. WO 86/03639 and that disclosure is incorporated herein by cross-reference.

Generally speaking, the invention provides an efficient method for the connectionless or connection oriented transfer of messages of arbitrary but finite length in a time multiplexed slotted environment with constrained destination resources.

The method provides for the efficient support of any type of addressing (short or extended, hierarchical or non-hierarchical) in the one environment, even in a 20 system with short slots.

In the one environment, the method can guarantee delivery of messages or provide a more efficient transfer at the expense of occasional message loss. In either case, there are minimal communication overheads and the utilisation of destination resources is maximised. Thus the method provides a wide range of options and considerable scope for achieving a range of performance-cost objectives.

BACKGROUND ART

In the data communications environment, information is generally exchanged in units called packets. These consist of an overhead necessary for the control and addressing of the unit through the data switch and 35 of the actual information. Typically the size of the information unit is not fixed but depends upon the message and the amount of information to be transferred.

Early packet switches handled the variable length packets as a whole unit, allocating all its communication resource to the transfer of the packet until its completion. There are a number of new packet switch designs emerging that switch only small fixed length slots. These switches are commonly referred to as Fast Packet Switches. Such switches are an improvement since they are generally simpler, may operate at higher speeds and allow for the support of real-time traffic.

If the Fast Packet Switch is to carry packet communications of variable length then it is necessary that the original packet message be segmented for transmission 50 over the switch and reassembled at the destination. The segmentation function is relatively simple only requiring that the message be divided into units of size equal to or smaller than the slot size. The transmission of the segments over the packet switch however requires 55 much more since it is necessary that the destination can receive and order all of the segments of the message. Thus it is required that there be a logical association between all of the slots of a single message. The reassembly function then reconstructs the original variable 60 length message from all of the received segments. Special care may need to be taken in the reassembly function to allow for the possibility that more than one message may need to be reassembled concurrently.

Some schemes have been developed to perform the 65 segmentation and reassembly function. However, these are limited either in the efficiency of the transport in the switch or in the performance of the reassembly func-

The problem with the slot overhead is compounded by addressing requirements. The common address field sizes used in data communications are 16 to 48 bits. With 48 bit addressing there is an overhead of 12 bytes per slot (source and destination address) in addition to the reassembly overheads. This approach to segmentation is clearly inefficient with small (less than 32 bytes) slot sizes.

40 BRIEF DESCRIPTION OF THE INVENTION

The invention also provides an apparatus for transmitting variable length messages on a network from a source to a destination in fixed length slots said apparatus including:

a segmentation machine for segmenting the message into fixed length slots which include a header field and a message segment, said machine including coding means for providing a source identifier field in the header of each slot, said source identifier field including a source identifier code which is uniquely associated with the message to be transmitted, and a reassembly machine located, in use, at the destination, said reassembly machine including control means for controlling reassembly of slots in accordance with the source identifier codes of the slots.

(c) Whether extended hierarchical or non hierarchical source and destination addressing is supported.

- (d) Whether the slot is carrying the first, last or continuation segment of a message.
- (e) Whether the message fits in a single segment.
- (f) Whether or not the source supports message transmission to multiple destinations concurrently.

The SI field 38 is a label which enables the logical association of all segments 40 belonging to the one message and hence enables them to be reassembled into the original message 20.

FIG. 2 diagrammatically illustrates a communications system comprising a source 42 which produces messages 20 of variable length for transmission on a network or fast packet switch 44 to a destination 46. The system includes a segmentation machine 48 coupled between the source 42 and the switch 44 and a reassembly machine 50 coupled between the switch 44 and the destination 46. The segmentation machine 48 converts messages 20 of indefinite length to slots 32 of fixed length for transmission on the switch 44. The reassembly machine 50 reassembles the slots 32 into the original message 20 for input to the destination 46. The segmentation and reassembly machines 48 and 50 would be located at respective nodes or access units coupled to the network.

FIG. 3 illustrates in more detail the use of the segmentation and reassembly machines 48 and 50 in a QPSX network of the type disclosed in WO 86/03639. The QPSX network comprises two unidirectional buses, bus A and bus B with data flowing in opposite directions, a central controller 2 and a number of distributed nodes or access units (AU's) 4 coupled between the buses A and B. Although each bus originates and terminates at the central controller 2 neither has a through connection, in normal circumstances. Each AU 4 has read taps 6 and 8 from the respective buses and lines 10 and 12 connected to unidirectional write couplers to the respective buses. The write transmit only in the direction of propagation of the respective buses. The read connections for each AU are attached to the bus ahead of the write connections and consequently the information read by each AU is unaffected by that written by it. In the illustrated arrangement, a source 42 is coupled to one of the access units 4 via the segmentation machine 48. The access unit transmits the message in fixed length slots on the network to the access unit 4 associated with the destination 46. Normally each access unit would have both segmentation and reassembly machines 48 and 50 to enable two way communications. The segmentation and reassembly machines 48 and 50 can be regarded as part of the interface IP 16.

FIG. 5 shows in more detail the connection of the segmentation and reassembly machines 48 and 50 to an access unit 4 of the type described in the aforementioned publication. FIG. 5 corresponds generally to FIG. 12 of that specification and hence need not be described in more detail here.

Each node in the network will have one or more unique SI's. Each SI can be used by the node for the transfer of a message. When the message transmission is complete the SI can be reused. Multiple SI's for a single node allow that node to transfer more than one message concurrently.

To describe the operation of the message transfer scheme, the segmentation of the message into slots is considered first and the action at the receiver is considered after that.

The train of slots 32 sent by the segmentation machine 48 is shown in FIG. 1. The first slot of a multiseg-

ment message will be identified as such by a COM (Beginning Of Message) code in the TYPE field 36. The SI field 36 is set by reference to the unique SI of the source node and the information field 40 contains the first segment of the message. Thus the DA field 22 of the message 20 is at the head of the information field 40. The following segments of the message until the last are each placed in the information fields of slots with the TYPE field 36 set to COM (Continuation Of Message) and the SI field 38 containing the source's code which is unique for this message. The last slot of a multisegment message has the TYPE field 36 coded as EOM (End Of Message), as shown.

For the transfer of a message 20 that only requires a single slot 32 the SSM (Single Slot Message) code is used in the TYPE field 36. The SI is not required in this case, however it is still used for consistency in operation.

An implementation of the segmentation machine 48 will now be described with reference to the state diagram which is shown in FIG. 6. In this diagram, the condition for a state transition is shown above the transition line and the action taken is below the line, in accordance with standard notation. This state machine will handle the receipt of a message from a single source at a time. If simultaneous receipt of more than one message is required, multiple state machines would be required.

To specify the communications between adjacent parts of the system, there are three communications primitives, as follows:

- (i) Request (REQ): This is a request to send a unit of data.
- (ii) Indication (IN): This is an indication that a unit of data has been received, and
- (iii) Confirmation (CONFIRM): This is a confirmation that a unit of data has been sent without error.

Between the machines 48 and 50 and the fast packet switch 44, (via the access units 4) communication primitives are prefixed by SAR_ACC hence there are three communication primitives as follows:

SAR_ACC Request
SAR_ACC Indication
SAR_ACC Confirm.

Also for source or destination equipment 42 and 46 such as a computer attached to the segmentation and reassembly machines 48 and 50, communication primitives are prefixed by APP_SAR hence there are three communication primitives as follows:

APP₁₃ SAR Request:
APP_SAR Indication
APP_SAR Confirmation

A segmentation machine operates on an unconfirmed APP SAR request, provided that sufficient resources are available. Tags for SAR_ACC requests and Source Identifiers (SIs) are common resources for all of these machines. Tags used in SAR_ACC requests are unique over all such requests from any segmentation machines. A tag is allocated on an SAR_ACC request and deallocated on an SAR_ACC confirm. The tag for the segments are local to the particular segment at the segmentation machine and is not transmitted on the network. Tags are reused when message transmissions have been completed. The tag is coded in TAG fields 52 and 54 in Request and Confirm primitives 56 and 58. The TAG fields 52 and 54 in the Request and Confirm primitives 56 and 58 are used for communications between the segmentation machine 48 and access unit 4, as

09919725 "077101

T_SEG: The T_SEG field in the SAR_ACC Indication indicates whether the slot 32 is the first segment of the message (BOM), a continuing segment (COM) or the final segment (EOM).

2.2 STATE R0: IDLE

R(00a) Idle—An SAR_ACC indication presents a segment with SSM set. A single segment message is assembled. The checksum is computed and if no error is indicated, an APP_SAR Indication presents the complete message to the destination 46.

R(01) Receiving—An SAR—ACC Indication presents a segment with BOM set and sufficient SAR resources are available to allow the reassembly machine to leave the idle state. Buffer allocated of size in octets equal to LENGTH indication. Resources are allocated and the data from the segment is buffered.

In FIG. 8, Action 1 beneath the transition line (11) includes the steps of storing the received segment 32 in sequence and computing the CHECKSUM over the message 20. Action 2 in transition lines R0(00a) and R1(10b) includes of reassembling the message 20, computing the CHECKSUM over the message 20. If CHECKSUM IN GOOD then APP_SAR_IND, otherwise discard the message.

In this state, the reassembly state machine is reassembling a message. The machine will remain in this state 45 until the complete message has been assembled or until the Reassembly Packet Timer PRT has expired or there is a problem with available resources. The RPT ensures that the reassembly machine is not effectively put out of service waiting for a lost EOM.

50 R(10a) Idle—When the Reassembly Packet Timer expires, the reassembly machine will return to the idle state, discarding the partially assembled message.

R(10b) Idle—When the complete message has been assembled (EOM received) the reassembly machine 55 computes the CHECKSUM. If the CHECKSUM indicates no error, an APP_SAR indication presents the received message to the destination 46. If the CHECKSUM indicates an error, the reassembly machine returns to the Idle state and releases resources.

60 R(11) Receiving—On each SAR_ACC indication with T_SEG=COM, the reassembly machine will buffer the segment in sequence.

The reassemble state machine diagram shown in FIG. 8 could be implemented in logic or by appropriate software.

FIG. 9 is another example of a state machine diagram for a reassembly machine which is particularly designed for use with the distributed queue system of the OPSX

This receive machine has two states: IDLE and WAIT. In the IDLE state the machine is not currently receiving any message. Thus in this state the machine will check for slots 32 with TYPE field 36 equal to BOM or SSM. In the case a BOM code is received the machine will check the DA in the information field. If the message is addressed to the station the machine enters the second state. In the case that SSM is detected in the IDLE state and the DA field matches, the length and information fields are copied and the AU 4 will indicate the higher layers that a message is received. The receive machine remains in the IDLE state after copying the slot. The higher layers refer to the higher layers in the Open System Interconnections (OSI) reference model.

The WAIT state is used to receive the slots 32 following the first of a multisegment message. In this state the slots with TYPE field 36 equal to COM and the SI 20 equal to that copied from the first slot of the message will be received. The information fields of these following slots are concatenated to form the complete message. New messages addressed to the given station in this state are ignored by the state machine. Further 25 receive machines are required if such messages are to be received. When the last slot of the message is detected by the receive state machine, the machine will copy the information field of the slot, indicating to the higher layers that a message is received and return to the 30 IDLE state. This completes the receipt of the message. To guard against the loss of the EOM slot, which would cause the receive state machine to be locked in the WAIT state, the RPT timer is used. This timer is started after each BOM or COM slot is received. If the timer 35 expires before the next slot is received, a failure in the transfer is assumed. The machine will then clear all copied slots and return to the IDLE state.

CONSTRAINED DESTINATION RESOURCES 40

In all practical implementations of destination facilities, i.e. resources such as buffers and processing capacity will be limited and loss of slots can occur unless a positive control mechanism is introduced to control communication between source and destination. 45

The method enables a source seeking use of the destination resources for reassembly of the original message to be temporarily held up until the required resources are available. The method implements a variety of access disciplines. Among others, the method supports 50 first come first served queueing for the destination resources described briefly below.

Control of access to destination resources is via a "ticket" handout scheme in which the "ticket number" establishes the position of source request in a distributed queue. The first encapsulated segment of a message sent by a source is considered by the destination as a request for resources. A "go ahead" message is returned to the source if the required resources are available, otherwise a "ticket" which uniquely identifies the sources position in a distributed queue of requests is returned. When resources become available the destination broadcasts the "ticket number" of the next source to be serviced. Under the assumption that resources are usually available, the delay incurred in waiting for the "go-ahead" is avoided by the source continuing to send segments of a message while it awaits the reply from the destination. In the case of a negative reply, that is a "ticket" is re-

—

3

10

30

55

65

To further describe the operation of the reassembly machine it is useful to consider the data flow paths through the machine when a slot is passed to the machine from the network. A fixed length slot 32 copied from the access unit 4 will enter both the DA Select circuit 70 and comparator block 72. The DA Select circuit 70 will temporarily store the full slot 32. Each of the sub-blocks 82 will only copy the SI field 38 of the slot.

If the segment received by the DA select circuit 70 is a BOM with a DA match then the circuit 70 will check if there are any free SI_Comp sub-blocks 82 available to control the receipt of the rest of the segments of the message. If not the segment is discarded. This event occurs when the reassembly machine is already receiving messages on all sub-blocks. This extra message is hence lost.

If there is an SI_Comp sub-block 82 available then the DA select circuit 70 will copy the SI from the segment and load it into the free SI_Comp sub-block 82, hence making that sub-block busy. The segment information field 40, that is the slot 32 minus the segment header fields 34, 36 and 38 is then passed via line 84 to the buffer selector circuit 70. The buffer selector circuit 76 in turn routes the segment 40 into the buffer 77 associated with the claimed SI_Comp sub-block 82. The association is signalled explicitly to the buffer selector circuit 74 from the SI_Comp block 82 via line 86.

For COM and EOM segments, that is the segments following the first of a message, the decision to copy these into the buffer is made by the SI_Comp circuit 72. Each sub-block 82 will compare the SI read from the incoming slot via line 88 with its own SI. If no sub-block has a match for the incoming SI then the segment is discarded. If there is a match, the information segment 40 is passed from the temporary storage in the DA select circuit 84 via line 84 to the sub-block 82. As with a BOM segment the buffer selector circuit 74 will route the segment to the buffer partition associated with the SI_Comp sub-block 82 that had a match for the SI. If the segment passed to the buffer partition is an EOM segment the reassembly of that message is complete. Hence the SI_Comp sub-block 82 is made free and the complete message passed out along line 80.

FIGS. 12a and b are flow charts which illustrate the logical steps involved in the reassemble machine 50, illustrated in FIG. 11.